The Evolving Design and Methods for Trials Evaluating the Safety of Candidate Vaginal Microbicides: A Systematic Review

Vicky Jespers, MD, PhD, Iona Y. Millwood, MD, PhD,† I. Mary Poynten, MBBS, MPH, PhD,‡ Lut Van Damme, MD, PhD,§ and John M. Kaldor, MD, PhD‡*

Abstract: Vaginal preexposure prophylaxis is a promising biomedical tool for HIV prevention. Although guidelines for the clinical assessment of microbicides are available, validated markers for product safety are lacking. To inform future microbicide and multipurpose vaginal product research, we reviewed the current and past safety methods used. We searched the Cochrane, EMBASE, and Ovid MEDLINE databases for clinical studies of vaginal products for the prevention of HIV that included safety evaluations. Ninety-seven clinical studies involving 21 products were identified: 63 lasted 14 days or less, 19 were longer in duration, and 15 were effectivess studies that included also safety as an outcome. Median sample size in the safety studies was 48 participants (range, 10–799). All studies reported on urogenital endpoint, 71% included colposcopy, and 67% assessed the vaginal microflora. Markers of vaginal epithelial inflammation, systemic absorption, and systemic toxicology assessments were evaluated in 29%, 26%, and 43% of studies, respectively. Excluding the effectiveness studies, these same assessments were done before 1998 in 33%, 7%, and 27% and after 2001 in 38%, 44%, and 60% of studies, respectively. Soluble inflammatory markers were introduced after 2001. Adverse event collection was reported in 73% of studies before 1998 and in 98% after 2001. In a previous review, we recommended that larger and longer safety studies were necessary to detect clinically important toxicities and to provide assurance that agents are ready for large-scale effectiveness trials. Here, we propose a stepwise clinical assessment that can be used for future guidance.

There has been an intensive search over more than 2 decades for a vaginally administered product, that is, a vaginal microbicide, that could prevent HIV acquisition when applied before exposure. After several products were found to be ineffective (Supplementary References http://links.lww.com/OLQ/A70 (S) [S1–5]), or possibly harmful [S6,7] in large-scale trials, the breakthrough for the field came in 2010 with the CAPRISA 004 trial reporting a 39% reduction in HIV acquisition, as well as a 51% reduction in herpes simplex virus type 2 acquisition, in women randomized to a regimen of 1% Tenofovir gel before and after sex [S8]. Although another trial of Tenofovir gel with a daily dosing regimen subsequently found no benefit, the CAPRISA 004 design is currently being repeated in South Africa, with high hopes of confirming the gel's effectiveness. Meanwhile, there are many other candidate products in various stages of development.[1] The lack of effect that emerged for several products in phase 3 trials raised questions about the design of phase I/II clinical trials. In a previous review, we reported that larger and longer safety studies were necessary to detect clinically important toxicities and to provide assurance that agents are ready for large-scale effectiveness trials.[2] To inform future safety microbicide and related mucosal vaginal product research,[3] we conducted a systematic review of the published literature. We aimed to identify emerging safety markers, to assess the use of the existing guidelines for the clinical development of microbicides,[4] and to develop up-to-date guidance for future research.

MATERIALS AND METHODS

We reviewed the following resources to obtain the names of vaginal microbicide products: International Microbicide Conferences 2006 to 2012,[5] Global Advocacy for HIV prevention database,[6] Alliance for Microbicide Development database,[7] and review papers on microbicide development and references within[8]. Using product names as keywords, a combined search of Cochrane Central Register of Controlled Trials (Third Quarter 2012), EMBASE (1980–2012), and Ovid MEDLINE (1950–2012) was conducted in September 2012 (week 36). Citations within publications were reviewed to identify further studies.

Clinical studies of candidate vaginal microbicides that described the methodology used for safety evaluation were included. Because the purpose of this review was to evaluate the methods used to assess safety in vaginal use for HIV prevention, studies conducted for a different indication (to assess spermicidal activity or contraceptive effectiveness, treatment of an existing or recurring condition, the prevention of mother to child transmission, or male studies) were not included. Gel distribution, acceptability, and feasibility studies were also not included. Publications identified in the search were reviewed by title, or title and abstract, by 2 reviewers. Those studies that potentially fulfilled the inclusion criteria were reviewed by full text. Studies were included based on agreement between both reviewers. Only studies published in English peer-reviewed journals were included.

The study design and methodology for safety assessments (Supplementary Table 1 http://links.lww.com/OLQ/A71) were extracted and tabulated by one reviewer and then checked by a second reviewer. To summarize design and methodology, studies were categorized into 3 groups: "short-term" safety studies

From the *Unit of Epidemiology and Control of HIV/STD, Institute of Tropical Medicine, Antwerp, Belgium; †Clinical Trial Service and Epidemiological Studies Units, University of Oxford, Oxford, United Kingdom; ‡Kirby Institute, University of New South Wales, Sydney, Australia; and §Bill & Melinda Gates Foundation, Seattle, WA

Conflicts of interest and source of funding: None of the following authors have an association that might pose a conflict of interest: Vicky Jespers, Iona Y. Millwood, I. Mary Poynten, Lut Van Damme, and John M. Kaldor. This work was supported by the European Commission (CHAARM 242135) and US Agency for International Development.

Correspondence: Vicky Achiel Jespers, MD, PhD, Unit of Epidemiology and Control of HIV/STD, Institute of Tropical Medicine, Antwerp, Belgium. E-mail: vjespers@itg.be.

Received for publication January 27, 2013, and accepted May 06, 2013.
Supplemental digital content is available for this article. Direct URL citations appear in the printed text, and links to the digital files are provided in the HTML text of this article on the journal's Web site (http://www.stdjournal.com).
DOI: 10.1097/01.olq.0000431070.38601.03

with 14 days or less of exposure to study products, "longer-term" studies with more than 14 days of exposure, and effectiveness studies. Assessments were reported as safety evaluations if they were conducted at least once before and once after exposure to study product and were described in the study methodology or results sections of the publication. To describe trends over time, studies were categorized into those conducted before 1995 and thereafter grouped by 3-to 4-year intervals (ie, 1987–1994, 1995–1997, 1998–2000, 2001–2004, 2005–2008). When the year the study was conducted was not reported, an estimate was made, assuming a period of 2 years between study conduct and publication of results.

RESULTS

Clinical studies fulfilling the inclusion criteria were identified for 21 of the 42 vaginal microbicide products that were identified (Supplementary Table 2 http://links.lww.com/OLQ/A72). Altogether, 97 studies instigated between 1987 and 2008 were included in the review (Table 1). The published studies included 29,473 women in 25 countries.

Study Designs

Sixty-three studies involved 14 days or less of exposure to study product. The randomized controlled trials (RCTs; 68%) had a variety of control groups: matched placebo [S9–25], universal placebo [S17,19,26–30], K-Y Jelly [S31–36], nonoxynol-9 (N–9) [S9,16,34,37,38], unspecified placebo [S39–42], methylcellulose placebo [S43,44], placebo intravaginal ring [S45], water [S46,47], and observation only [S44]. Seven (11%) studies had 100 or more participants, and 252 HIV-infected women were enrolled in 8 studies. Sexually abstinent women were enrolled in 35 (55%) studies, sexually active women were enrolled in 15 (24%) studies, and 13 (21%) studies had a subset of each. Nineteen longer-term safety studies involving more than 14 days of study product exposure were identified for 10 products. Four of the studies involved a total of 156 HIV-infected women [S48–51]. All studies enrolled sexually active women, and 1 also included sexually abstinent women. The frequency of exposure to study product ranged from at least one application per week to twice a day or more if coital frequency was higher. Fifteen studies evaluating the effectiveness of a vaginal microbicide reported on safety measures. Most were RCTs with a placebo control arm. There were 2 (13%) exceptions, both early N-9 studies (1988–1989), with one study enrolling HIV-serodiscordant couples and using a condom only arm [S52,53]. Frequency of exposure to study products was coitally dependent in 14 studies and once daily in one study [S54] (Table 1).

In most studies, the product was a gel formulation. Studies of N-9 and Praneem Polyherbal also involved cream, film, suppository, or tablet formulations, and 1 trial of lime juice involved the application of a soaked tampon. Some studies of Acidform, Buffergel, and cellulose sulfate used the microbicide gel vaginally in conjunction with either a diaphragm [S36,51,55] or the Duet barrier device [S56]. Dapivirine ring studies tested the reservoir-type ring in 3 studies [S45,57] and the matrix ring in 1 study [S57]. Most safety trials were implemented in populations with low HIV incidence. The effectiveness studies were all conducted in sub-Saharan Africa, with other study sites in Asia [S7,58] and the United States [S3].

Safety Evaluations

Nearly all studies reported the evaluation of urogenital symptoms, except for 1 short-term [S59] and 1 effectiveness study [S60]. Pelvic examinations for genital signs were routinely conducted in all safety studies and in all but 1 of the effectiveness studies (where it was only performed if "clinically indicated," based on prespecified criteria [S2]). Colposcopic examination was performed for all participants in 50 (79%) short-term safety, 16 (84%) longer-term safety, and 3 (20%) effectiveness studies. Of the studies that used colposcopy, 63% described the use of a standardized methodology, predominantly the World Health Organization/CONRAD 2000 or 2004 guidelines.[9,10] Assessment of adverse events (AEs), other than urogenital events, was reported in 58 (92%) short-term safety, 18 (95%) longer-term safety, and 10 (67%) effectiveness studies. In the remaining studies, the study reports did not indicate whether or not AEs were routinely collected (Fig. 1).

Systemic toxicology parameters were evaluated in 31 (49%) short-term and 9 (47%) longer-term safety studies. In these studies, hematology results were always reported, along with 1 or more of coagulation tests, electrolytes, and renal and hepatic function tests. Plasma drug levels were tested in a total of 25 (26%) studies, including 4 PRO2000 studies, 3 studies of SPL7013, 1 study each of Savvy and Terameprocol, and in all short-term and longer-term safety studies of antiretrovirals (ARVs), including Dapivirine, Tenofovir, and UC781. In recent studies, ARV levels were also measured in vaginal lavage and/or swab samples [S26–29,45,61–64] and in vaginal biopsies [S29,62,64].

Alterations in the vaginal microflora were assessed by testing for bacterial vaginosis (BV) in 39 (64%) short-term safety, 14 (74%) longer-term safety, and 5 (33%) effectiveness studies. Culture was performed in 18 (19%) studies, most of which were short term. Vaginal candidiasis was assessed in 34 (54%) short-term, 14 (74%) longer-term, and 9 (60%) effectiveness studies.

Twenty-four (38%) short-term safety, 3 (16%) longer-term safety, and 1 (7%) [S54] effectiveness study assessed vaginal epithelial inflammation. Leukocytes were counted in cervicovaginal lavage samples, cervical or vaginal smear specimens, or on Gram stains in 21 studies (22%). Four studies (4%), all conducted in the UK, collected biopsy specimens for histological analysis [S18,23–25]. Thirteen studies involved an analysis of soluble cytokines in cervicovaginal lavage and/or swab samples. The most commonly measured markers were interleukin (IL)-1α, IL-1β, IL-6, and IL-8.

Trends Over Time for Short and Longer Safety Studies

For 22 (27%) of the 82 studies, the period of conduct was not reported, so an estimate was made. Before 1995, only 9 studies had been published, including 4 effectiveness studies. Thereafter, there was an average of 3 to 5 safety studies per year, with numbers increasing from 2001 onward. The proportion of trials designed as RCTs over time remained consistent at 70%. The median number of participants decreased from 125 in the earlier years to 36 in more recent years. The frequency and duration of exposure did not vary with time (Table 2).

Testing for markers of vaginal epithelial inflammation, systemic absorption, and systemic toxicology increased from 33%, 7%, and 27% of studies before 1998 to 38%, 44%, and 60% after 2001, respectively. The first reported systemic absorption study involved specimens from a 1997 PRO2000 study [S25]. The laboratory testing was performed considerably later on the stored samples. Assessment of soluble inflammatory markers was introduced after 2001. Before 1998, inflammation was assessed by histology or by detection of semen as a proxy of inflammation or by measurement of leukocytes in vaginal fluid.

TABLE 1. Summary of Studies Included in the Review

	Published Studies, *n*	Randomized Controlled Studies, *n*	No. Participants		Frequency of Exposure, doses/day		Duration of Exposure, day		Supplementary References*†
			Median	Range	Median	Range	Median	Range	
Short-term safety studies (1–14 d)									
Acidform	3	2	20	18–81	1	1	6	4–14	[S36,66–68]‡
Buffergel	3	1	81	30–98	1	1–2	14	2–14	[S36,56,68,69]‡
Carraguard	4	2	47	25–60	1	1–1	7	7–14	[S43,44,59,70]
Cellulose acetate phthalate	1	0	5	—	1	—	14	—	[S71]
Cellulose sulfate	5	5	59	48–180	2	1–4	14	6–14	[S31–35]
Dextrin sulfate	2	2	68	36–100	2	1–3	9	5–14	[S18,21,23]
Docusate sodium	1	1	13	—	1	—	7	—	[S21]
Invisible condom	2	1	131	41–221	1–2	—	14	—	[S17,72]
Lime juice	2	2	36	25–47	1–2	1–2	13	12–14	[S46,47]
Menfegol	1	1	125	—	0.5–8§	—	14	—	[S11]
N-9	9	5	48	19–534	1	1–4	14	1–14	[S21,24,39,40,42,73–78]
Polystyrene sulfonate	1	1	49	—	1	—	6	—	[S16]
Praneem polyherbal	3	1	23	20–58	1	1–1	10	7–14	[S41,79,80]
PRO2000	6	3	33	10–73	1	1–2	14	1–14	[S13,14,25,81–84]
Savvy	3	3	64	60–105	1	1–2	7	3–14	[S9,37,38]
Tenofovir	3	1	49	30–84	1–2	1–2	14	—	[S28,64,85]
Terameprocol	1	1	14	—	1	—	7	—	[S15]
Dapivirine (TMC120)	5	4	18	12–64	1	1–2	7	7–11	[S12,29,45,62]
UC781	3	3	48	45–60	2	1–2	1	1–14	[S26,27,30]
Universal placebo	1	1	30	—	2	—	14	—	[S22]
SPL7013 (Vivagel)	4	3	45	12–61	2	1–2	10	5–14	[S10,19,20,63,86]
Subtotal 1	63	43	48	10–534	1	1–8	14	1–14	
Longer-term safety studies (>14 d)									
Acidform	1	1	120	—	1¶	—	182	—	[S55]
Buffergel	2	1	413	27–799	1¶–2	—	59	28–91	[S3,65,87]‡**
Carraguard	3	3	165	55–400	1	1–7 a week	365	182–365	[S88–92]
Cellulose sulfate	1	1	119	—	1¶	—	182	—	[S51]
Dextrin sulfate	2	2	93	77–109	2	1–3	28	28	[S48,93]
Invisible condom	1	1	194	—	2	—	63	—	[S94]
N-9	3	2	170	20–320††	1¶	—	30	28–182	[S50,95,96]
Praneem polyherbal	1	1	100	—	1¶	—	182	—	[S97]
PRO2000	2	2	490	180–799	1¶–2	—	59	28–91	[S3,49]‡**
Dapivirine	3	3	36	24–112	2	—	42	28–42	[S57,61]
Subtotal 2	19	17	119	20–799	1	1–3	63	28–365	
Effectiveness studies					**				
Buffergel	1	1	3100	—	1	—	365	—	[S3]‡**
Carraguard	1	1	6202	—	1	—	495	270–720	[S5]
Cellulose sulfate	2	2	1516	1398–1644	1	—	365	—	[S6,7]
N-9	6	4	276	110–1294	1	—	365	350–425	[S52–54,58,60]
PRO2000	2	2	6242	3100–9385	1		365	365–365	[S3,4]‡**
Savvy	2	2	2148	2142–2153	1	—	365	365–365	[S1,2]
Tenofovir	1	1	1085	—	2‡‡	—	365	365	[S8]
Subtotal 3	15	13	1398	110–9385	—	—	365	42–720	
Total	97	71	60	10–9385	—	—	14	1–720	

*There may be more than 1 publication resulting from a study and one publication can describe 2 or 3 studies.
†The references are provided as supplementary material in the Supplementary Reference List http://links.lww.com/OLQ/A70. Numbers are preceded with an S throughout the text.
‡This study involved 2 products, each of which is counted under each product heading.
§Application every other day, once, twice, 4, or 8 times daily.
¶Application before coitus.
**This study is a combination of a phase II long-term safety study and a phase III effectiveness study and is counted under both headings.
††One study sample size not reported.
‡‡One application before and after sex.

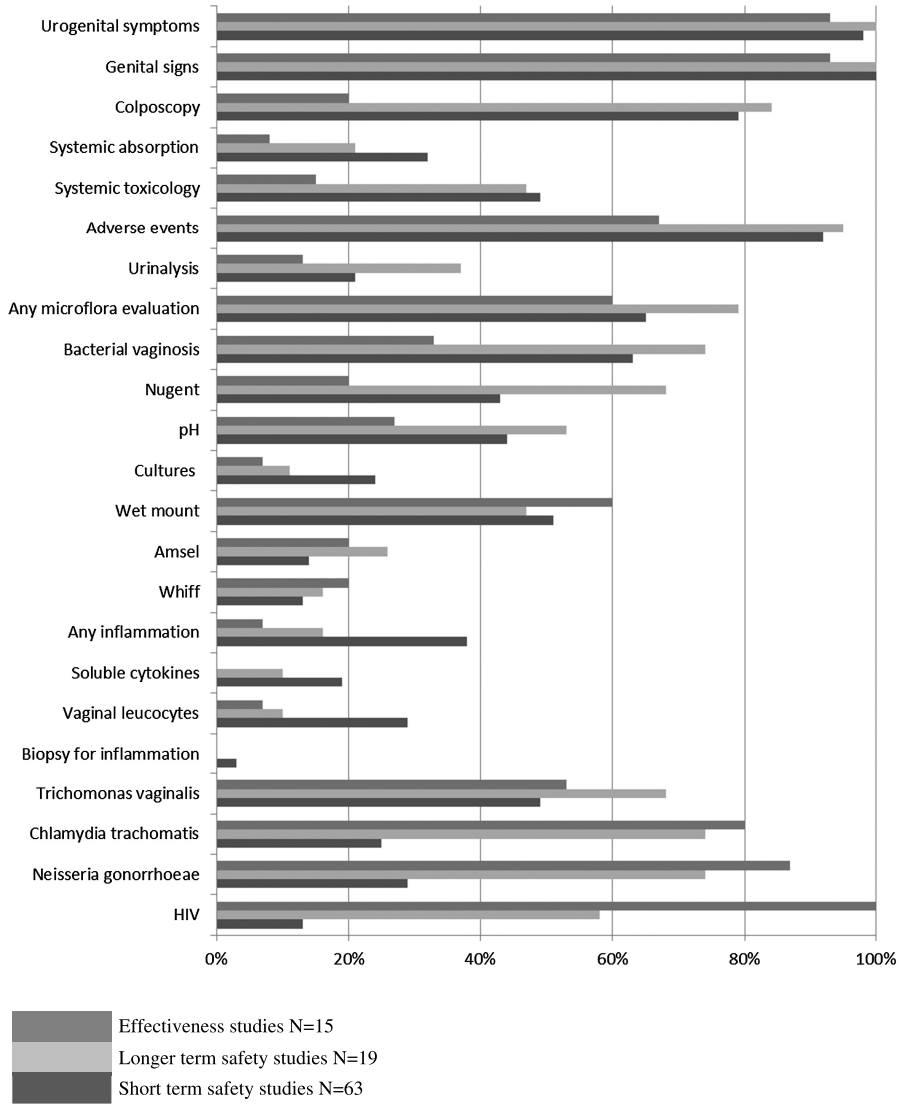

Figure 1. Summary of safety evaluations[a]. Effectiveness studies (n = 15). Longer-term safety studies (n = 19). Short-term safety studies (n = 63). [a]Evaluations conducted at least once before and once after exposure to study product. [b]Vaginal microflora: defined by any evaluation including BV as diagnosed by clinical or laboratory diagnosis, Nugent score,[16] pH, vaginal cultures, wet mount, or Amsel score.[30]

Adverse event collection was done in 73% of studies before 1998 and in 98% after 2001. There was an increase in the proportion of studies evaluating the microflora from 2001 onward.

DISCUSSION

This systematic review of 97 microbicide clinical studies indicates that the methods used in the assessment of safety have been very diverse and have changed markedly over the past 10 years. Almost two thirds of published studies involved product exposure of 2 weeks or less. Safety studies tended to be smaller in later years, with the median number of participants decreasing from 125 in 1994 to 36 in 2008. These changes may be a result of recommendations from the International Working Group on Microbicides in the early 2000s[4] after consultation on the negative results of the N-9 trial and the cellulose sulfate trial being stopped early because of potential harm [S6,7]. They also may reflect the approach of initiating a number of pilot safety studies to explore safety in great detail, including the effect of semen, before embarking on larger trials.

Clinical Examination and Guidelines

The evaluation of urogenital symptoms, with assessment of local irritation and tissue damage by genital symptoms and clinical examination including colposcopy, was almost universal. This demonstrates that most of the studies reviewed here generally conform to the International Working Group on Microbicides guidelines[11] and the guidelines by Mauck et al.,[4] describing study design and safety endpoint considerations for phase I to III clinical studies. Although the value of colposcopic findings in microbicide trials has been questioned because of poor correlation of colposcopic AE findings with histological

TABLE 2. Trends in Safety Design Over Time for Short- and Longer-Term Safety Studies

	Year the Study Was Started									
	1987–1994		1995–1997		1998–2000		2001–2004		2005–2008	
No. safety studies	5		10		15		27		25	
Short term	5		7		11		21		19	
Longer term	0		3		4		6		6	
Study design	**n**	**%**	**n**	**%**	**n**	**%**	**n**	**%**	**n**	**%**
Randomized controlled trial	4	80	7	70	8	57	22	81	17	68
Sample size description										
Formal hypothesis	2	40	2	20	4	27	4	15	8	35
Exploratory/feasibility	2	40	0	0	2	13	13	48	6	26
No explanation	1	20	8	80	9	60	10	37	9	39
	Median	**Range**	**Median**	**Range**	**Median**	**Range**	**Median**	**Range**	**Median**	**Range**
Participant no.	125	20–534	38	13–320	70	18–400	57	12–180	36	5–799
Frequency of exposure	1[*†]	0.5–8	1	1–2	1[‡]	1–4	1[§¶]	1–3	1**	1–4
Duration of exposure	14	14–14	10	1–182	14	3–365	14	1–182	14	1–91
Safety evaluations	**n**	**%**	**n**	**%**	**n**	**%**	**n**	**%**	**n**	**%**
Urogenital symptoms	5	100	9	90	15	100	27	100	25	100
Genital signs	5	100	10	100	15	100	27	100	25	100
Colposcopy done	4	80	10	100	10	67	24	89	18	72
Guidelines specified	1		8		9		20		9	
Guidelines not specified	3		2		1		4		9	
Systemic absorption	0	0	1	10	1	7	8	30	15	60
Systemic toxicology	1	20	3	30	5	33	12	44	19	76
AEs	3	60	8	80	14	100	26	96	25	100
Any vaginal microflora evaluation[††]	2	40	7	70	9	60	23	85	15	65
BV[††]	2		6		9		22		14	
Nugent score	0		6		5		18		11	
pH	0		6		4		16		12	
Cultures with H_2O_2	0		7		0		4		6	
Wet mount	1		4		7		18		11	
Any inflammation evaluation[§§]	0	0	5	50	2	13	14	52	6	22
Soluble cytokines	0		0		0		8		5	
Vaginal leukocytes	0		4		3		9		4	
Biopsy	0		4		0		0		0	
Semen	0		3		0		1		2	
Sexually transmitted infection	**n**	**%**	**n**	**%**	**n**	**%**	**n**	**%**	**n**	**%**
Trichomoniasis	2		6		8		18		10	
Chlamydia	1		6		6		11		6	
Gonorrhoeae	2		6		6		12		6	
HIV	0		2		2		7		8	

*One study application every other day, once, twice, 4, or 8 times daily.
†One study application every other day, once, twice, or 4 times daily.
‡One study 3 or more applications per week, 1 study application once daily plus before coitus; 1 study application twice daily plus before coitus maximum 3 times daily; 1 study application before coitus; 1 study at least 3 times a week; 1 study involved application twice or 4 times daily.
§One study involved application once or twice daily.
¶One study application twice daily or before coitus maximum 3 times daily.
**Two studies once or twice daily; one study once, twice, or thrice daily.
††Vaginal microflora: defined by any evaluation including BV, Nugent score, pH, vaginal cultures, wet mount, or Amsel score.
‡‡Diagnosed by clinical or laboratory diagnosis.
§§More than one method was used in several studies.
H_2O_2 indicates hydrogen peroxide.

evaluations[12] [S24, 25] and reports of poor colposcopic interobserver agreement[13] [S23], this remained a standard feature of safety studies. A recent reanalysis of data from 9 N-9 studies found that although more abnormal findings were seen with colposcopy, there was no benefit greater than that of naked eye examination.[14] The authors recommended naked eye examinations in microbicide studies and to only consider colposcopy for pilot and early product safety studies.

It is important that systemic drug levels and toxicity are evaluated for all products.[4] Absorption was measured for 8 of the 21 products in the current review, and toxicity was not assessed in most safety studies. In the case of ARVs, multiple studies to obtain sufficient pharmacokinetic data on local exposure, local distribution, and persistence of product are needed to define optimum product dosage and formulation.

Evolving Design and Emerging Safety Methods

The range and type of safety evaluations, particularly for nonclinical examination markers (ie, toxicology, absorption, microflora evaluation, and inflammation methods), were not standardized across studies. Several new safety methods have emerged since the publication of the guidance documents.

Because an unbalanced vaginal microflora has been associated with an increased risk of HIV,[15] it is essential to

investigate the effect of topical products on vaginal microflora. This was studied for virtually all products (except polystyrene sulfonate and Terameprocol); however, the methods used for microflora evaluation have not been comprehensive. Only half the safety studies performed a Gram stain for the identification of BV by Nugent score,[16] and less than a quarter investigated the effects of the microbicide on hydrogen peroxide–producing lactobacilli with culture methods [S65]. This low number likely reflects the labor-intense culture methods and the need to incubate a fresh sample. Culture methods have a bias toward nonfastidious bacteria, and resulting data will, as such, be restricted. To overcome the difficulties of culture methods, DNA-based methods (amplification of the *16SrDNA* genes) have recently been used in microbicide trials for the identification of the bacteria present in the vagina.[17–19] A quantitative polymerase chain reaction method can quantify a selection of species representing a "healthy" and a "disturbed flora" before, during, and after product use.[20] This method has been applied in the vaginal safety evaluation of Polymun antibodies in a gel[17] and in the IPM013 Dapivirine vaginal ring study[17] (as yet unpublished). This quantitative method is recommended for future safety studies.

Inflammatory reactions in vaginal tissues can increase the risk of HIV transmission or acquisition[21] through cell activation, cell inhibition, cell injury, and influx of target cells. Therefore, it is critical to ensure that locally applied products do not cause inflammation. In early trials, inflammation was identified by microscopy on vaginal biopsies and detection of semen and leukocytes in vaginal fluid. Measurement of the effect of microbicides on cytokines, chemokines, innate mediators, and target cells in vaginal specimens was rare in early years, but reports on these factors substantially increased from 2001 onward[22] [S14]. Since then, IL-8 and the ratio of IL-1RA to IL-1 have been validated and proven valuable as independent biomarkers of the vaginal immune balance[23] [S35,47]. Functional ex vivo models studying vaginal fluid before and after product exposure and assessing intrinsic antiviral (HIV and herpes simplex virus type 2) and antibacterial (*Escherichia coli* and *Staphylococcus aureus*) activity in vitro [S28] were originally designed to test the efficacy of a microbicide [S13]. However, these models proved to be good surrogate measurements of endogenous antimicrobial activity and thus microbicide safety. By triangulating the ex vivo antimicrobial activity with the soluble antimicrobial markers (eg, lactoferrin and SLPI) and the soluble inflammatory (eg, IL-1 and IL-6) and anti-inflammatory (eg, IL-1ra) markers, a more in-depth safety picture of the product on the immune balance is provided. This approach has been used by Keller et al.[24] in a PRO2000 and Tenofovir phase I study and in a recent contraceptive gel study with Acidform [S14,28]. The polarized ex vivo ectocervical model studies the effect of a microbicide and/or its formulation on the epithelial integrity of the explant. This model, validated against N-9 as a positive control for the study of Tenofovir gel, revealed the negative aspects of the hyperosmolar nature of the formulation.[25] Finally, sexually transmitted infections may act as confounders by affecting inflammatory parameters.[4] Sexually transmitted infection testing was understandably low in short safety studies conducted in very-low-risk settings. More testing was performed in the longer safety and effectiveness studies enrolling sexually active participants from countries with high sexually transmitted infection prevalence.

Trial Design

The proportion of studies describing the sample size calculation with or without a formal hypothesis was low. The median sample size for the safety studies was 54 participants. Studies of this size have sufficient power to detect differences between a placebo and product group for a frequently present (>70%) continuous variable. For example, to have 90% power of detecting an effect of a product reducing the mean log count of *Lactobacillus crispatus* by 1.5, the number of participants required in each arm is 38. However, this sample size is insufficient to detect an increase from 10% to 20% in the frequency of a dichotomous endpoint such as epithelial disruption. In addition, the sample sizes and trial duration have generally not been adequate to assure safety in long-term, high-frequency users.[2] We recommend that all studies are powered to answer a specific hypothesis and that power calculations are described in the articles.[26] Furthermore, the VOICE trial showed that adherence to gel was low despite women reporting high usage.[27] A better understanding and ways of improving adherence are needed to ensure that the product under study is actually used and safety can be adequately assessed.

Although most studies reviewed were randomized, many of the studies did not have an adequate control group. Given that the power to detect differences in outcomes was invariably low, the use of a comparison group could assist in interpretation of results. Several studies used an N-9 product or matched placebo as a control. Given the adverse safety profile of N-9 and some other excipients, these products would not now be considered an appropriate comparison unless used as a positive control. The universal placebo gel has been developed to provide a more standardized comparison and is widely used in current studies [S22].

Reporting of Results of Safety Studies

A limitation of this review is that our report does not reflect all completed studies because there is a delay between conducting a study and publishing the results. It is also possible that safety evaluations were conducted but not mentioned in the methodology or results of articles. For example, the routine collection of AEs was not reported in several studies, in particular those conducted in earlier years. This could be caused by a focus on local safety, a nonsystematic approach to recording of AEs, or the style of reporting those studies. Some bias in reporting is also a possibility, with a particular test or methodology only being reported, and sometimes only in the results section, if noteworthy results were found.

Strategy for Microbicide Clinical Safety Assessment

On the basis of our review findings, as well as recent technological and methodological advances, we propose a stepwise standardized clinical assessment of the safety of vaginal products (Table 3). Pilot human studies should thoroughly investigate any changes in the vaginal compartment before initiating larger-scale trials. Early phase I and rectal ancillary substudies can be run in parallel. Large phase I studies should be adequately powered to detect statistically meaningful differences between product and control arms.[2] Postcoital studies are needed to address the question of continued product activity in the presence of semen.

Given the extensive resources required for long-term trials, there should be a more judicious selection of products before large-scale clinical testing.[28] A cautionary approach should be taken in the safety assessment of potential microbicides, to provide as near an assurance of safety as possible and increase the confidence among communities participating in these important HIV prevention trials.

TABLE 3. Endpoints and Stepwise Plan for Microbicide Clinical Safety Trial Design

	Design	Population	Product Use	Endpoints on Clinical Examination	Endpoints Other
First clinical study*	Open label first few women to protect participants' safety followed by RCT	Not sexually active Low risk	1–7 d	AE reporting GU symptoms GU signs including colposcopy	Systemic toxicology A selection of the markers listed below
Pilot study*	RCT 20 women Rectal assessment	Not sexually active Low risk	14 d	AE reporting GU symptoms GU signs	Systemic absorption. Systemic toxicology. Microbiome: pH, Gram stain for Nugent scoring, cultures, species PCR†, pyrosequencing.† Inflammation: vaginal leukocytes; biopsy cervical/rectal: epithelial integrity and permeability; target cells†: phenotype and activation status; soluble factors in vaginal fluid: proinflammatory cytokines (IL-1β, IL-6, IL-1α), anti-inflammatory cytokine IL-1ra, α-chemokines (IP-10, IL-8), antimicrobial proteins† (elafin, defensin, SLPI), growth factors: G-CSF; (β-chemokines MIP-1β); endocervical cells†: phenotype and activation status target cells. Ex vivo functional models: antiviral and antibacterial activity of lavage fluid†
Early phase I*	RCT Dosing schedules Crossover design 20 women	Sexually active Low risk	30 d	AE reporting GU symptoms GU signs	A selection of the above markers informed by translation and preclinical study data
Late phase I	RCT	Adolescents/pregnant women	1–30 d	AE reporting GU symptoms GU signs	A selection of the listed markers
Large phase I	RCT Postcoital substudy	Sexually active Low risk	30 d and more	AE reporting GU symptoms	A selection of the listed markers informed by translation and preclinical study data
	200 women			GU signs	Functional models: antiviral and antibacterial activity of vaginal fluid in the presence of semen†
Phase II	RCT 200 women	Target population	6 mo	AE reporting GU symptoms GU signs	A small selection of the listed markers

*These studies can be combined.
†Emerging markers.
PCR indicates polymerase chain reaction; d, days; mo, months; AE, adverse event; GU, genitourinary; IP-10, Interferon gamma-induced protein 10; SLPI, Secretory Leukocyte Peptidase Inhibitor; G-CSF, granulocyte-colony stimulating factor; MIP-1β, macrophage inflammatory protein 1 beta.

REFERENCES

1. Ariën K, Jespers V, Vanham G. HIV sexual transmission and microbicides. Rev Med Virol 2011. E-pub ahead of print.
2. Poynten IM, Millwood IY, Falster MO, et al. The safety of candidate vaginal microbicides since nonoxynol-9: A systematic review of published studies. AIDS 2009; 23:1245–1254.
3. Thurman AR, Doncel GF. Herpes simplex virus and HIV: genital infection synergy and novel approaches to dual prevention. Int J STD AIDS 2012; 23:613–619.
4. Mauck C, Rosenberg Z, Van Damme L. Recommendations for the clinical development of topical microbicides: An update. AIDS 2001; 15:857–868.
5. Microbicides Conferences. Available at: http://www.global-campaign.org/Int-microbicides-Conferences.htm. Accessed January 26; 2013.
6. Global Advocacy for Prevention (AVAC). Available at: http://data.avac.org/Timeline aspx. Accessed January 26; 2013.
7. Alliance for Microbicide Development database. 2012. Available at: http://www.microbicide.org/. Accessed December 14; 2012.
8. Balzarini J, Van Damme L. Microbicide drug candidates to prevent HIV infection. Lancet 2007; 369:787–797.
9. CONRAD/WHO. Manual for the standardization of colposcopy for the evaluation of vaginal products. Update 2000.
10. CONRAD/WHO. Manual for the standardization of colposcopy for the evaluation of vaginal products. Update 2004.
11. International Working Group on Vaginal Microbicides. Recommendations for the development of vaginal microbicides. AIDS 1996; 10:1–6.
12. Mauck C, Baker J, Birnkrant D, et al. The use of colposcopy in assessing vaginal irritation in research. AIDS 2000; 14:2221–2227.
13. Guido RS, Cardinal-Busse BJ, McIntyre-Seltman K, et al. Colposcopically obtained images of vulvar, vaginal, and cervical epithelium for assessment of safety of intravaginal agents among normal adolescent females: Comparison with in vivo exam and interobserver agreement. J Low Genit Tract Dis 2003; 7:259–263.
14. Mauck CK, Weiner DH, Lai JJ, et al. Colposcopy: Still useful in microbicide safety trials? Sex Transm Dis 2012; 39:465–469.
15. Taha TE, Hoover DR, Dallabetta GA, et al. Bacterial vaginosis and disturbances of vaginal flora: Association with increased acquisition of HIV. AIDS 1998; 12:1699–1706.
16. Nugent R, Krohn M, Hillier S. Reliability of diagnosing bacterial vaginosis is improved by a standardized method of gram stain interpretation. J Clin Microbiol 1991; 29:297–301.
17. Jespers V, Morris G, Nel A, et al. Vaginal microbiota assessment by real time PCR in the Mabgel1 study and the IPM013 Dapivirine ring trial. Microbicides Conference, Sydney, Australia Oral presentation: Abstract 508. 2012.
18. Verhelst R, Verstraelen H, Claeys G, et al. Cloning of 16S rRNA genes amplified from normal and disturbed vaginal microflora suggests a strong association between *Atopobium vaginae*, *Gardnerella vaginalis* and bacterial vaginosis. BMC Microbiol 2004; 4:16.
19. Ravel J, Gajer P, Fu L, et al. Twice-daily application of HIV microbicides alter the vaginal microbiota. MBio 2012; 3.
20. Jespers V, Menten J, Smet H, et al. Quantification of bacterial species of the vaginal microbiome in different groups of women, using nucleic acid amplification tests. BMC Microbiol 2012; 12:83.
21. Gumbi PP, Nkwanyana NN, Bere A, et al. Impact of mucosal inflammation on cervical human immunodeficiency virus (HIV-1)–specific CD8 T-cell responses in the female genital tract during chronic HIV infection. J Virol 2008; 82:8529–8536.
22. Fichorova RN. Guiding the vaginal microbicide trials with biomarkers of inflammation. J Acquir Immune Defic Syndr 2004; 37(suppl 3): S184–S193.
23. Fichorova RN, Lai JJ, Schwartz JL, et al. Baseline variation and associations between subject characteristics and five cytokine biomarkers of vaginal safety among healthy non-pregnant women in microbicide trials. Cytokine 2011; 55:134–140.
24. Keller MJ, Carpenter CA, Lo Y, et al. Phase I randomized safety study of twice daily dosing of Acidform vaginal gel: Candidate antimicrobial contraceptive. PLoS One 2012; 7:e46901.
25. Rohan LC, Moncla BJ, Kunjara Na Ayudhya RP, et al. In vitro and ex vivo testing of tenofovir shows it is effective as an HIV-1 microbicide. PLoS One 2010; 5:e9310.
26. Turner L, Shamseer L, Altman DG, et al. Consolidated standards of reporting trials (CONSORT) and the completeness of reporting of randomised controlled trials (RCTs) published in medical journals. Cochrane Database Syst Rev 2012; 11:MR000030.
28. Marrazzo J, Ramjee G, Nair G, et al. Pre-exposure prophylaxis for HIV in women. Daily Oral Tenofovir, Oral Tenofovir/Emtricitabine, or Vaginal Tenofovir Gel in the VOICE Study (MTN 003). 20th Conference on Retroviruses and Opportunistic Infections, Atlanta, US. Abstract 26LB. 2013.
28. Niedbala RS, Mauck C, Harrison P, et al. Biomarker discovery: Validation and decision-making in product development. Sex Transm Dis 2009; 36(3 suppl):S76–S80.
29. Adverse Event Data NIH Division of AIDS Table for Grading the Severity of Adverse Events. Available at: http://rsc.tech-res com/Document/safetyandpharmacovigilance/gradingtables aspx. Accessed January 26, 2013.
30. Amsel R, Totten PA, Spiegel CA, et al. Nonspecific vaginitis. Diagnostic criteria and microbial and epidemiologic associations. Am J Med 1983; 74:14–22.